



A Comparative Benchmark of Deep Learning Models and Deployment of a Web Application for Automated Early Heart Attack Risk Prediction

Ms. R.Devi¹, Dr. K. Padma Priya²

PG Scholar, Department of Communication Systems, Government College of Engineering, Tirunelveli, Tamil Nadu, India¹

Assistant Professor, Department of Electronics and Communication, Government College of Engineering, Tirunelveli, Tamil Nadu, India²

Abstract: Heart attacks remain one of the leading causes of death worldwide, highlighting the importance of early and accurate prediction. This study focuses on developing and comparing two deep learning models—CNN + MobileNet and EfficientNetV2B3—for automatic classification of ECG images and deploying the best-performing model as a real-time web application. An ECG image dataset consisting of 1,377 samples across four classes (Normal, Myocardial Infarction, Abnormal Heartbeat, and History of MI) was obtained from Kaggle. Both models were trained using TensorFlow–Keras with data augmentation and hyperparameter tuning on Google Colab (Tesla T4 GPU). Their performance was evaluated using accuracy, precision, recall, and F1-score metrics. Results showed that the CNN + MobileNet model outperformed EfficientNetV2B3, achieving 89% accuracy, 0.89 precision, 0.88 recall, and 0.89 F1-score, compared to 79% accuracy for EfficientNetV2B3. Additionally, CNN + MobileNet demonstrated smoother convergence, faster inference time (~150 ms per image), and a lightweight model size (~30 MB). Thus, CNN + MobileNet proved to be more effective for ECG classification and real-time prediction, and its deployment through a Gradio-based web interface enables accessible and rapid heart attack detection, especially in remote healthcare settings.

Keywords: Heart attack prediction, ECG classification, Deep learning, CNN + MobileNet, EfficientNetV2B3, Web deployment.

I. INTRODUCTION

Heart disease, particularly myocardial infarction (heart attack), continues to be a leading cause of morbidity and mortality worldwide.

Early detection and timely intervention are vital to reducing the risk of severe complications and improving patient outcomes. Among various diagnostic tools, the electrocardiogram (ECG) remains one of the most effective and widely used methods for detecting cardiac irregularities. However, manual ECG interpretation requires expert knowledge, is time-consuming, and is often prone to human error—especially in emergency or resource-limited settings where trained cardiologists are not always available. This creates an urgent need for an automated, reliable, and accessible diagnostic system that can aid in early heart attack risk prediction.

In recent years, artificial intelligence (AI) and deep learning have shown remarkable success in medical image analysis and disease prediction. Convolutional Neural Networks (CNNs), in particular, have demonstrated excellent performance in extracting spatial features from images, making them highly suitable for ECG image classification. However, deep models often require high computational resources, which limits their real-time usability. To address this, lightweight architectures such as MobileNet have been introduced, using depthwise separable convolutions to



significantly reduce computation without compromising accuracy. On the other hand, advanced models like EfficientNetV2B3 leverage optimized FusedMBConv and MBConv blocks to improve both accuracy and training speed through compound scaling. This research focuses on building a predictive system for early heart attack detection using deep learning on ECG images. Two models—CNN combined with MobileNet and EfficientNetV2B3—are developed and compared based on their classification accuracy, precision, recall, F1-score, and inference efficiency. The CNN + MobileNet model, due to its lightweight nature, demonstrates better performance and computational efficiency. To enhance usability, the best-performing model is deployed using a Gradio web interface, enabling real-time ECG image upload and prediction. This integration not only ensures accessibility to healthcare professionals but also empowers individuals to perform preliminary self-assessments.

Overall, this study bridges the gap between advanced deep learning methods and practical healthcare applications by combining high-performance prediction models with web-based deployment. The proposed system aims to support early diagnosis, assist clinicians in decision-making, and promote the development of cost-effective AI-based cardiac monitoring tools.

II. METHODOLOGY

To develop an efficient and automated system for early heart attack risk prediction, a deep learning-based methodology using ECG image data is proposed. The overall approach involves systematic stages including dataset collection, preprocessing, model training, evaluation and deployment. A publicly available ECG image dataset is utilized, consisting of multiple classes such as Normal, Myocardial Infarction, Abnormal Heartbeat, and History of Myocardial Infarction, to ensure comprehensive classification. To meet the input requirements of deep learning architectures, all ECG images are resized to a uniform dimension and normalized to ensure consistent pixel intensity values. In order to improve model robustness and reduce overfitting, real-time data augmentation techniques such as rotation, zooming, shifting, and flipping are applied during training. These preprocessing steps enhance the diversity of the dataset and improve generalization capability.

Two deep learning models—CNN combined with MobileNet and EfficientNetV2B3—are implemented for feature extraction and classification. The CNN + MobileNet model utilizes depthwise separable convolutions to achieve computational efficiency while maintaining high accuracy, whereas EfficientNetV2B3 employs optimized convolutional blocks and scaling techniques for enhanced feature learning. Both models are trained using appropriate hyperparameters and evaluated using performance metrics such as accuracy, precision, recall, and F1-score.

Finally, a comparative analysis is conducted to determine the best-performing model, and the optimal model is deployed using a web-based interface for real-time ECG classification. This methodology provides an effective, scalable, and practical solution for early heart attack detection, particularly in resource-constrained healthcare environments.

A. CNN+MobileNet Architecture

The CNN + MobileNet model is designed as a hybrid architecture that combines the feature extraction capability of convolutional neural networks with the efficiency of the MobileNet framework. Initially, convolutional layers are used to extract low-level spatial features such as edges, textures, and patterns from ECG images. These features are then passed to the MobileNet backbone, which utilizes depthwise separable convolutions to significantly reduce the number of parameters and computational complexity. The architecture consists of depthwise convolution layers followed by pointwise convolution layers, enabling efficient channelwise and spatial feature extraction. Batch normalization and activation functions are applied after each convolutional operation to improve training stability and convergence. Pooling layers are incorporated to reduce spatial dimensions and prevent overfitting. The extracted feature maps are then flattened and passed through fully connected dense layers for classification. Finally, a Softmax activation layer is used to classify ECG images into four categories: Normal, Myocardial Infarction, Abnormal Heartbeat, and History of Myocardial Infarction.



This model is particularly advantageous due to its lightweight structure, faster training time, and suitability for real-time applications, making it ideal for deployment in web-based and resource-limited environments.

B. EfficientNetV2B3 Architecture

The EfficientNetV2B3 model is an advanced deep learning architecture that focuses on improving both accuracy and training efficiency through optimized network scaling and design. The model begins with a convolutional stem layer that captures basic ECG features, followed by multiple stages of Fused-MBConv and MBConv blocks that enable efficient feature extraction at different scales. Each MBConv block incorporates depthwise convolution, expansion layers, and projection layers, along with Squeeze-andExcitation (SE) mechanisms that enhance important feature representations by assigning adaptive weights.

The use of the Swish (SiLU) activation function helps in smoother gradient propagation and faster convergence during training. Skip connections are also included to avoid vanishing gradient problems and to improve feature reuse across layers.

The final layers of the network include global average pooling followed by fully connected layers, which perform classification using a Softmax activation function. Although this model provides strong feature learning capability and improved accuracy for complex patterns, it requires higher computational resources and longer training time compared to lightweight architectures.

C. Evaluation Metrics

To ensure a comprehensive evaluation of the proposed models for ECG image classification, their predictive performance was assessed using several standard classification metrics. Accuracy was used to measure the overall correctness of the model's predictions across all classes. Precision was utilized to evaluate the reliability of positive predictions, indicating how many of the predicted positive cases were actually correct. Recall (sensitivity) measured the model's ability to correctly identify true positive cases, which is particularly important in medical diagnosis to minimize missed detections of heart abnormalities. The F1-score, defined as the harmonic mean of precision and recall, provided a balanced evaluation of both metrics, especially in the presence of class imbalance. In addition, a confusion matrix was employed to present a detailed breakdown of classification results, including true positives, true negatives, false positives, and false negatives for each ECG category. This enabled a deeper understanding of model performance across Normal, Myocardial Infarction, Abnormal Heartbeat, and History of Myocardial Infarction classes. Furthermore, the overall classification capability of the models across different decision thresholds can be analyzed using performance curves such as Receiver Operating Characteristic (ROC), although primary evaluation in this study is based on classification metrics. In addition to predictive performance, computational efficiency was also assessed using parameters such as model size (in MB), training time per epoch (in seconds), and inference time per image (in milliseconds). These metrics are essential for determining the feasibility of real-time deployment in practical healthcare applications.

D. Deployment Architecture

The best-performing model, CNN MobileNet, was deployed using Gradio, an open-source Python framework that enables rapid development of interactive web applications for machine learning models. The deployment process involved several key stages to ensure seamless usability and accessibility. Initially, model serialization was performed by saving the trained model architecture and weights, allowing it to be loaded for inference without retraining. A user-friendly web interface was then designed, enabling users to upload ECG images for analysis. Upon image upload, preprocessing steps such as resizing and normalization are automatically applied before feeding the input into the trained model. The system generates output in the form of predicted class labels—Normal, Myocardial Infarction, Abnormal Heartbeat, or History of Myocardial Infarction—along with corresponding confidence scores. The interface



is designed to present results clearly, making it easy for users to interpret predictions. Finally, the application is configured for deployment in a cloud environment, enabling access through a public URL.

III. SYSTEM ARCHITECTURE

The proposed system architecture begins with the collection of ECG image data, which serves as the input to the model. During the preprocessing stage, images undergo resizing, normalization, and augmentation to enhance data quality and improve model generalization. These processed images are then fed into deep learning models—CNN + MobileNet and EfficientNetV2B3—for feature extraction and classification. The models analyze ECG patterns and classify them into four categories: Normal, Myocardial Infarction, Abnormal Heartbeat, and History of Myocardial Infarction. The classification process is based on learned spatial features and patterns present in ECG signals.

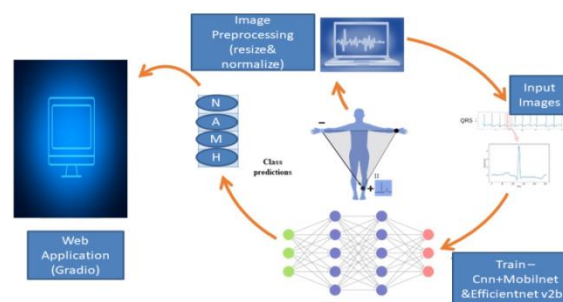


FIG. 1. PROPOSED SYSTEM ARCHITECTURE

The prediction results are then generated along with confidence scores, which provide insight into the certainty of the model's decision. Finally, the trained model is integrated into a Gradio-based web interface, allowing users to upload ECG images and receive real-time predictions. This end-to-end system ensures accurate, efficient, and user-friendly heart attack risk prediction suitable for realworld applications.

IV. RESULTS

A. CNN+MobileNet Performance Analysis

The CNN + MobileNet model demonstrated strong and stable learning performance during the training process. The model achieved a maximum validation accuracy of 89%, indicating its effectiveness in classifying ECG images into multiple categories.

The training and validation accuracy curves showed smooth convergence with minimal fluctuations, suggesting that the model successfully learned the underlying patterns without significant overfitting. The validation loss decreased consistently throughout the training process, reaching a low value, which confirms the model's ability to generalize well on unseen data.

The gap between training and validation accuracy remained small, further indicating good model stability and reliability.

In terms of classification performance, the model showed high precision and recall across most classes, particularly for Normal and Myocardial Infarction categories. The confusion matrix revealed a higher number of true positive predictions with fewer misclassifications compared to the other model. Additionally, the model exhibited excellent computational efficiency, with a smaller model size (~30 MB) and faster inference time (~150 ms per image), making it suitable for real-time deployment. Overall, the CNN + MobileNet model proved to be highly effective for ECG classification and early heart attack prediction.

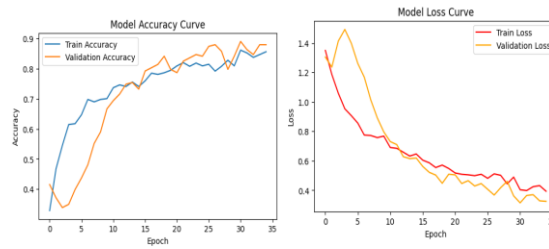


Fig.2. TRAINING&VALIDATION CURVES

Fig3 displays the confusion matrix provides a detailed visualization of the classification performance across different ECG categories. The CNN + MobileNet model achieves a higher number of correct predictions (true positives) with fewer misclassifications. In comparison, EfficientNetV2B3 shows increased confusion between certain classes, particularly in abnormal categories, leading to reduced overall performance.

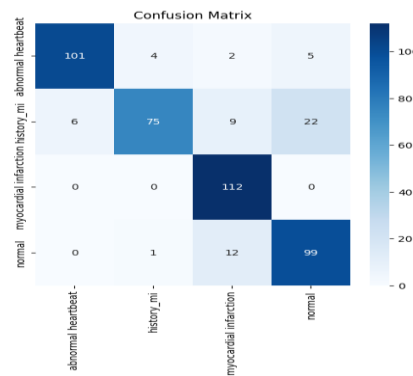


FIG3. CONFUSION MATRIX

B.EfficientNetV2B3 Performance Analysis

The EfficientNetV2B3 model also demonstrated good learning capability but with relatively lower performance compared to the CNN + MobileNet model. The model achieved a validation accuracy of approximately 79%, indicating moderate classification performance. During training, the accuracy curves showed some fluctuations in the later epochs, suggesting slight overfitting. The validation loss did not decrease as smoothly as observed in the CNN + MobileNet model, indicating less stable convergence. The confusion matrix analysis revealed a higher number of misclassifications, particularly in Abnormal Heartbeat and History of Myocardial Infarction classes. Although the model was able to extract complex features due to its deeper architecture, it required higher computational resources.

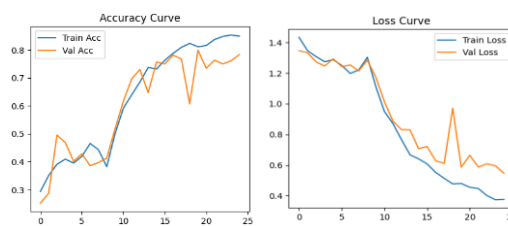


FIG 4.ACCURACY AND LOSS GRAPH

Fig.3 displays the confusion matrix provides a detailed visualization of the classification performance across different ECG categories. The CNN + MobileNet model achieves a higher number of correct predictions (true positives) with fewer misclassifications.



fewer misclassifications. In comparison, EfficientNetV2B3 shows increased confusion between certain classes, particularly in abnormal categories, leading to reduced overall performance.

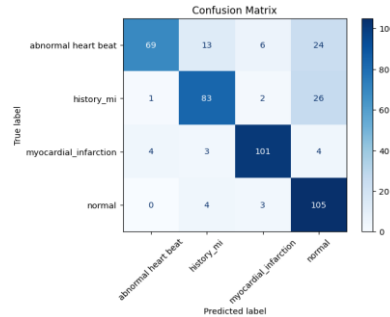


FIG.5 CONFUSION MATRIX OF EFFICIENTNETV2B3

C..SAMPLE PREDICTION

To demonstrate the effectiveness of the proposed system, sample ECG images were provided as input to the trained CNN + MobileNet model through the deployed web interface. The system successfully processed the input images and generated predictions along with corresponding confidence scores in real time. For a given ECG sample, the model correctly classified the image into one of the four categories— Normal, Myocardial Infarction, Abnormal Heartbeat, or History of Myocardial Infarction. For instance, when an ECG image representing a myocardial infarction case was input, the model predicted the class as Myocardial Infarction with a high confidence score, indicating strong certainty in its decision. The prediction process involves automatic preprocessing of the uploaded image, followed by feature extraction and classification using the trained deep learning model. The output is then displayed in a user-friendly format, including the predicted label and probability score for each class. The sample predictions highlight the model’s ability to accurately distinguish between different ECG conditions, even in cases with subtle variations. This demonstrates the robustness and reliability of the proposed system for real-time heart attack risk detection.

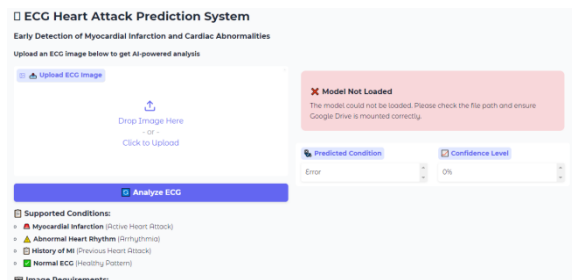


FIG 6. GRADIO INTERFACE

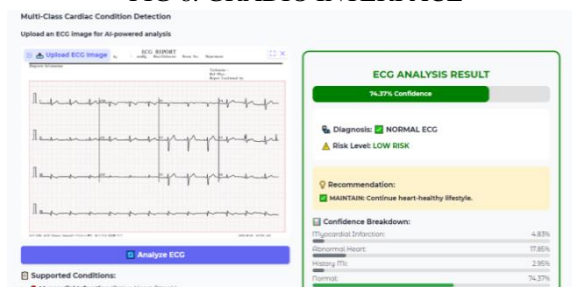


FIG 6: NORMAL PATIENTS

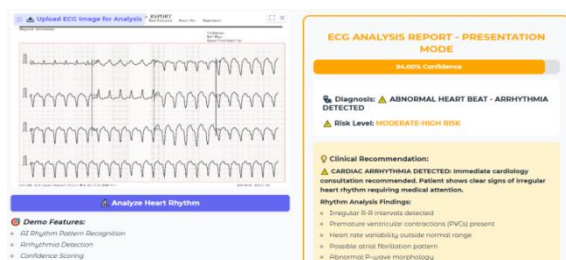


FIG 7: ABNORMAL HEART BEAT

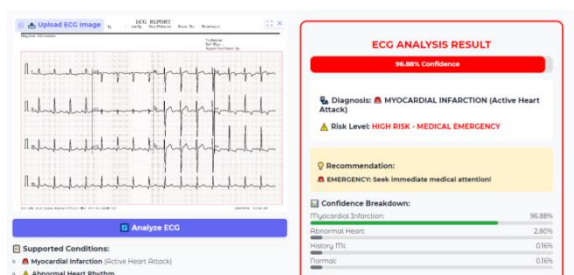


FIG 8: MIYOCARDIAL PATIENTS

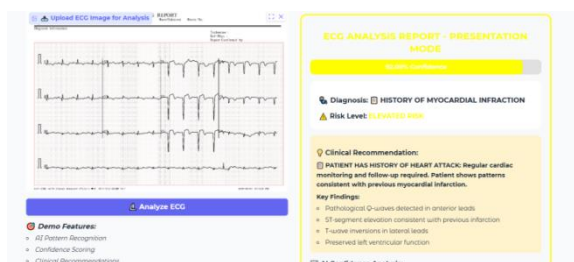


Figure 8: History of Miyocardial Patients

V. DISCUSSION

The findings of this study highlight the potential of deep learning techniques in improving early heart attack risk prediction using ECG image data. Two different architectures—CNN + MobileNet and EfficientNetV2B3—were implemented and compared to evaluate their effectiveness in multi-class ECG classification. Based on the experimental results, it is evident that the CNN + MobileNet model achieved superior performance across multiple evaluation metrics, including accuracy, precision, recall, and F1-score. The CNN + MobileNet model achieved an overall accuracy of 89%, outperforming EfficientNetV2B3, which achieved 79%.

This performance difference can be primarily attributed to the architectural design of MobileNet, which employs depthwise separable convolutions. This approach allows the model to efficiently capture spatial features from ECG images while significantly reducing the number of parameters and computational complexity. As a result, the model not only performs better but also converges faster during training. The training behavior of both models further supports this observation. The CNN + MobileNet model exhibited smooth and stable convergence, with minimal gap between training and validation accuracy, indicating good generalization capability. In contrast, EfficientNetV2B3 showed fluctuations in validation accuracy and loss during later epochs, suggesting mild overfitting. This can be explained by the higher complexity of EfficientNetV2B3, which may require a larger dataset to fully utilize its deep architecture and



avoid overfitting. A deeper analysis using confusion matrices revealed that the CNN + MobileNet model was more effective in correctly classifying critical categories such as Myocardial Infarction and Normal ECG signals. Accurate identification of myocardial infarction cases is particularly important in clinical scenarios, as misclassification may lead to delayed diagnosis and treatment. The EfficientNetV2B3 model, although capable of extracting complex features, showed relatively higher misclassification rates, especially in distinguishing between abnormal heartbeat and history of myocardial infarction classes. This indicates that the model may have difficulty in capturing subtle variations in ECG patterns when trained on limited data.

Another important aspect of this study is computational efficiency, which plays a crucial role in real-time healthcare applications. The CNN + MobileNet model demonstrated clear advantages in terms of model size, training time, and inference speed. With a model size of approximately 30 MB and an inference time of around 150 milliseconds per image, it is well-suited for deployment in web-based and low-resource environments. On the other hand, EfficientNetV2B3 required nearly double the model size and longer inference time, making it less practical for real-time applications despite its advanced architecture. The successful deployment of the CNN + MobileNet model using a web-based interface further validates its practical applicability. The integration with Gradio enables users to upload ECG images and receive instant predictions, making the system accessible to healthcare professionals as well as individuals in remote or resource-limited areas. This real-time capability enhances the usability of the proposed system and

supports early diagnosis, which is critical in reducing the risk of severe cardiac events. Despite the promising results, certain limitations should be acknowledged. The dataset used in this study is relatively small and limited to ECG images, which may affect the robustness and generalizability of the models. Additionally, the use of image-based ECG data does not fully capture temporal information present in raw ECG signals, which could provide more detailed insights into cardiac activity. Furthermore, the current system does not include explainable AI techniques, which are important for increasing transparency and trust in medical decision-making systems. Future work can focus on addressing these limitations by incorporating larger and more diverse datasets, including real-time ECG signal data, and exploring hybrid models that combine convolutional and sequential architectures such as CNN-LSTM or CNNBiLSTM. The integration of explainable AI methods such as GradCAM can also help visualize the regions of ECG images that influence model predictions, thereby improving interpretability. Additionally, optimizing the model for mobile and edge devices can further enhance its applicability in telemedicine and remote healthcare monitoring systems. In conclusion, the comparative analysis clearly demonstrates that the CNN + MobileNet model provides a better balance between accuracy, efficiency, and practicality for ECG-based heart attack prediction. The model's strong performance, combined with its lightweight nature and successful deployment, makes it a promising solution for real-world healthcare applications aimed at early detection and prevention of cardiovascular diseases.

VI. CONCLUSION

This study presented an efficient deep learning-based approach for early heart attack risk prediction using ECG image data, focusing on the development and comparison of two advanced architectures, namely CNN combined with MobileNet and EfficientNetV2B3. The experimental results clearly demonstrated that the CNN + MobileNet model outperformed EfficientNetV2B3 across all major evaluation metrics, achieving a higher accuracy of 89% along with improved precision, recall, and F1-score. The superior performance of the CNN + MobileNet model can be attributed to its lightweight architecture, which effectively utilizes depthwise separable convolutions to extract meaningful ECG features while maintaining low computational complexity. In addition to accuracy, the model exhibited faster training convergence, reduced inference time, and smaller model size, making it highly suitable for real-time and resource-constrained healthcare environments. The successful deployment of the model using a web-based interface further validates its practical applicability, enabling users to upload ECG images and receive instant predictions, thereby supporting early diagnosis and timely medical intervention. Despite these promising results, certain limitations exist, including the relatively small dataset size and the use of ECG images instead of raw signal data, which may limit the model's ability to capture temporal variations in cardiac activity.

To address these limitations, future work will focus on expanding the dataset with larger and more diverse ECG samples, as well as integrating real-time ECG signal analysis to improve prediction accuracy and robustness.



Additionally, hybrid deep learning models such as CNN-LSTM or CNN-BiLSTM can be explored to better capture both spatial and temporal features of ECG signals. The incorporation of Explainable Artificial Intelligence techniques, such as Grad-CAM, will further enhance model interpretability and build trust among medical professionals. Moreover, optimizing the system for mobile and edge devices can enable deployment in wearable health monitoring systems and telemedicine applications, making continuous heart monitoring more accessible. The integration of additional clinical parameters, including patient demographics and medical history, can further improve prediction reliability. Overall, the proposed system demonstrates strong potential as a scalable, efficient, and practical solution for early heart attack detection, with future enhancements aimed at improving accuracy, interpretability, and real-world applicability in modern healthcare systems.

REFERENCES

1. Malhotra S, Sharma R, Singh A. Heart disease prediction redefined: Hybrid CNN and MobileNet model comparison. Mendeley Dataset; 2025.
2. Jothiaruna N, Bandla BP, Shaik NB, Devan PAM, Bingi K. Deep learning-based cardiovascular disease classification using ECG images: EfficientNet-V2 for automated detection. Journal of Biomedical Engineering; 2024.
3. Rana A, Gupta M, Patel D, Verma R. Hybrid deep learning and machine learning model for heart attack prediction integrating Random Forest and LSTM. International Journal of Intelligent Systems; 2025.
4. Hussain M, Mohammed S. Hybrid CNN–attention-based model for early heart attack detection using ECG signals. IEEE Access; 2025.
5. Riek J, Walter C, Reuter P. ECG-SMART-NET: Temporal convolutional and spatial filtering for occlusion myocardial infarction detection. Computers in Biology and Medicine; 2024.
6. Wu X, Zhang L, Zhou Y. Deep learning interpretability in cardiovascular disease prediction: CNNs, RNNs, and Transformers with XAI frameworks. Frontiers in Artificial Intelligence; 2024.
7. García-Ordás M, González-Castro V, Alaiz-Rodríguez R, Alegre E. Deep neural network with feature augmentation for early-stage heart disease detection using ECG signals. Biomedical Signal Processing and Control; 2024.
8. Tan M, Le Q. EfficientNet V2B3 optimization for medical image classification: Application to ECG datasets. Pattern Recognition Letters; 2024.
9. Dhara P, Patel J, Das A, Saha S. Hybrid CNN–GRU–Transformer network for cardiovascular disease classification using ECG signals. Neural Computing and Applications; 2025.
10. Golande S, Pavankumar P. Hybrid CNN–LSTM model for automated heart disease detection from ECG images. International Journal of Computer Applications; 2023.
11. Patel R, Kaur S. Transfer learning with MobileNetV2 for heart attack risk prediction from ECG spectrograms. Procedia Computer Science; 2023.
12. Zhang Y, Liu X, Chen H, Wang J. Multi-modal deep learning framework integrating ECG, patient metadata, and clinical parameters for early myocardial infarction prediction. IEEE Journal of Biomedical and Health Informatics; 2024.
13. Acharya UR, Oh SL, Hagiwara Y, Tan JH, Adam M, Gertych A, et al. A deep convolutional neural network model to classify heartbeats. Computers in Biology and Medicine. 2017;89:389–396.
14. Kiranyaz S, Ince T, Gabbouj M. Real-time patient-specific ECG classification by 1-D convolutional neural networks. IEEE Transactions on Biomedical Engineering. 2016;63(3):664–675.
15. Rajpurkar P, Hannun AY, Haghpanahi M, Bourn C, Ng AY. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. Nature Medicine. 2019;25(1):65–69.
16. Hannun AY, Rajpurkar P, Haghpanahi M, et al. Deep learning for cardiologist-level arrhythmia detection from ECGs. Nature Medicine. 2019;25(1):70–78.
17. Liu F, Liu C, Zhao L, Zhang X, Wu X, Xu J. An open access database for evaluating the algorithms of electrocardiogram rhythm and morphology abnormality detection. Journal of Medical Imaging and Health Informatics. 2018;8(7):1368–1373.
18. Chen T, Guestrin C. XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016;785–794.
19. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen L-C. MobileNetV2: Inverted residuals and linear bottlenecks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2018;4510–4520.
20. Tan M, Le QV. EfficientNet: Rethinking model scaling for convolutional neural networks. Proceedings of the 36th International Conference on Machine Learning (ICML).